



University of Tennessee, Knoxville
**Trace: Tennessee Research and Creative
Exchange**

DataONE Sociocultural and Usability &
Assessment Working Groups

Communication and Information

6-24-2010

Baseline assessment of scientists' data sharing practices.

Carol Tenopir

University of Tennessee - Knoxville, ctenopir@utk.edu

Follow this and additional works at: https://trace.tennessee.edu/utk_dataone



Part of the [Library and Information Science Commons](#)

Recommended Citation

Tenopir, Carol, "Baseline assessment of scientists' data sharing practices." (2010). *DataONE Sociocultural and Usability & Assessment Working Groups*.

https://trace.tennessee.edu/utk_dataone/68

This Presentation is brought to you for free and open access by the Communication and Information at Trace: Tennessee Research and Creative Exchange. It has been accepted for inclusion in DataONE Sociocultural and Usability & Assessment Working Groups by an authorized administrator of Trace: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

Baseline Assessment of Scientists' Data Sharing Practices

Carol Tenopir, University of Tennessee

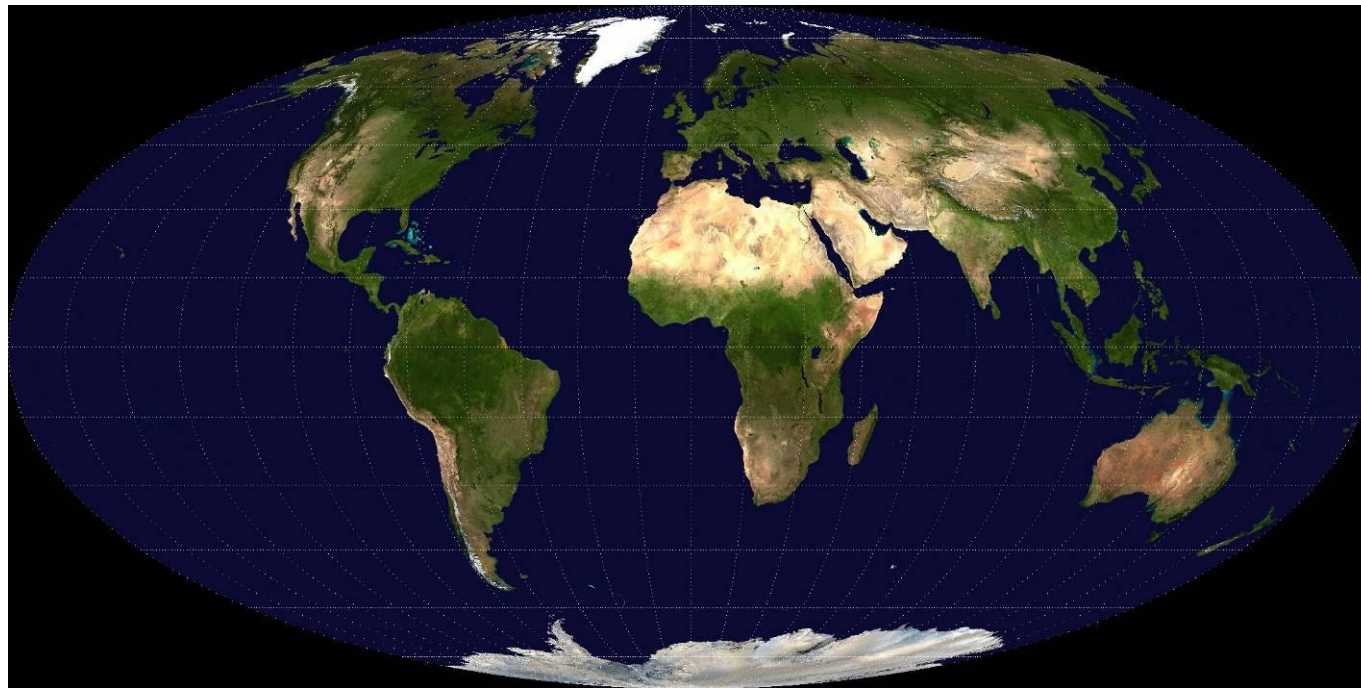
NSF DataNet program will build new types of organizations that will...

- integrate library and archival sciences, cyberinfrastructure, computer & information sciences, and domain science expertise to:
 - provide reliable digital preservation, access, integration, and analysis capabilities for science and/or engineering data over a decades-long timeline

DataONE

(Data Observation Network for Earth)

P.I., Bill Michener, University Libraries, Univ. New Mexico



- Environmental science challenges
- Cyberinfrastructure challenges
- DataONE: A solution
 - Building on existing CI
 - Creating new CI
 - Changing science culture and institutions

- Libraries & digital libraries
- Academic institutions
- Research networks
- NSF- and government-funded synthesis & supercomputer centers/networks
- Governmental organizations
- International organizations
- Data and metadata archives
- Professional societies
- NGOs
- Commercial sector



To measure the current state of data needs, practices, knowledge of standards, and motivations regarding data collection, access, and preservation

GOOD PRACTICES

TIME

Computer – IT Personnel



Scientists



Public Officials



**Citizen-scientists
Students & Teachers**

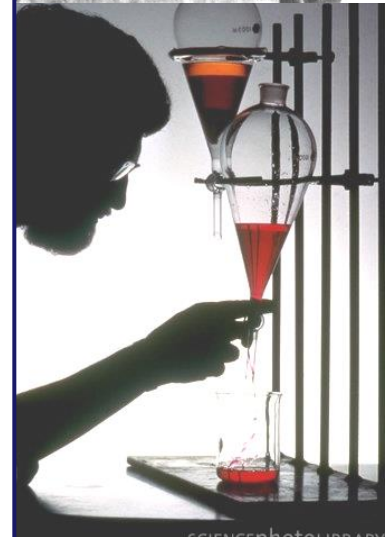


**Libraries
Librarians**

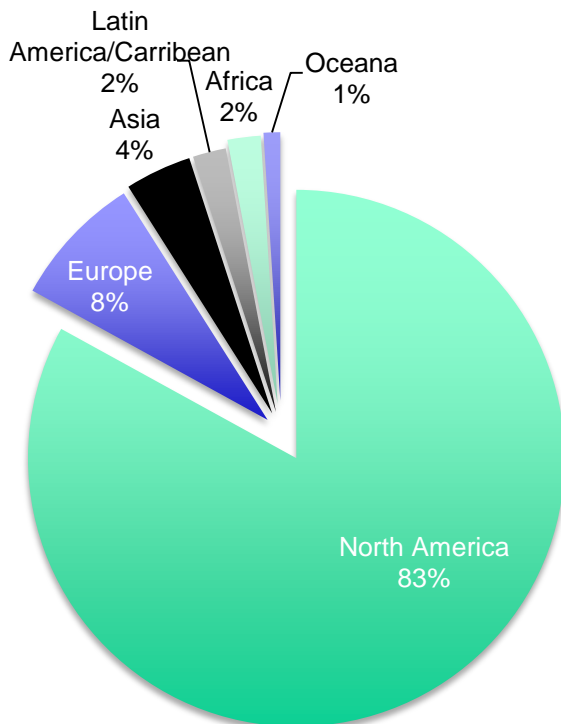
Baseline Assessment of Scientists: distribution and responses



- **Scientists - various work sectors**
- **Via champions**
- **As of June 2010 N=1000**
- **Preliminary results N=923**

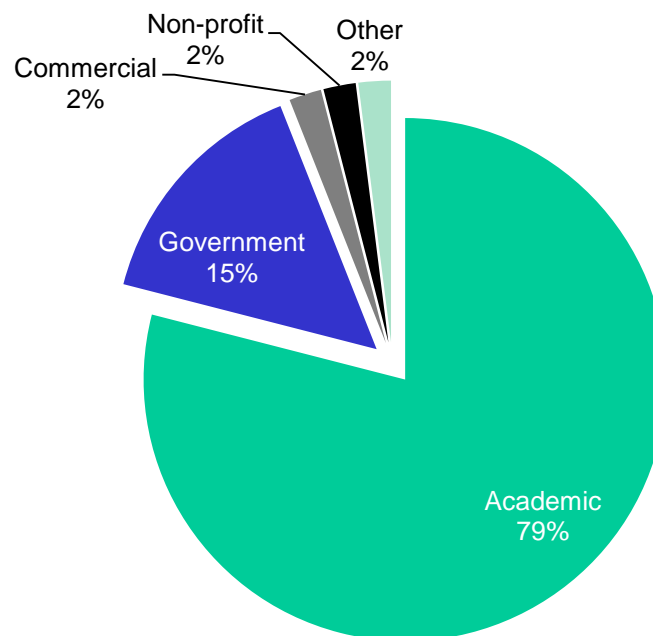


Place of Employment by World Region

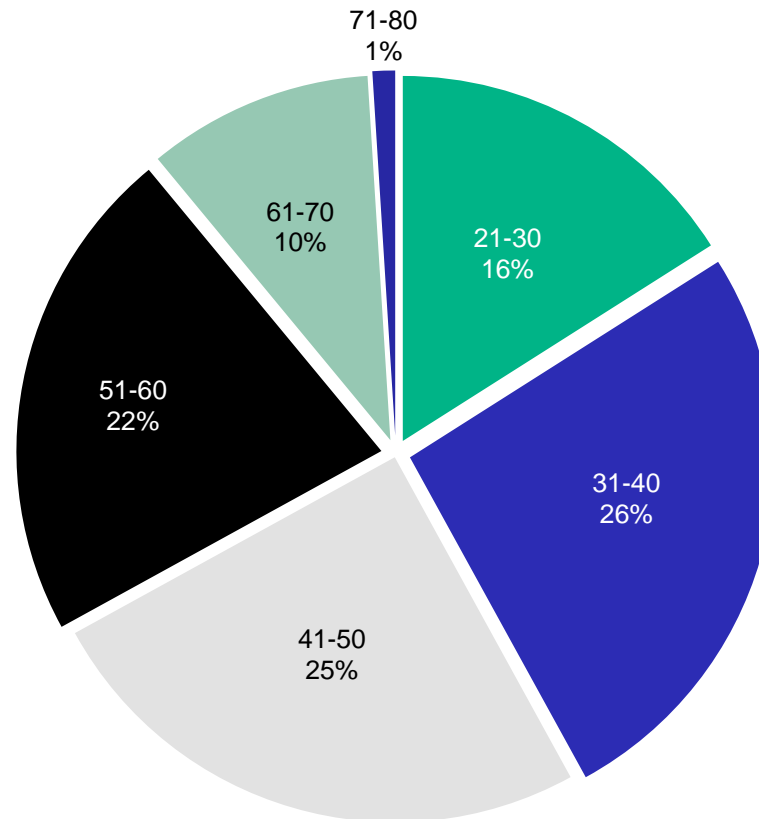


N=909

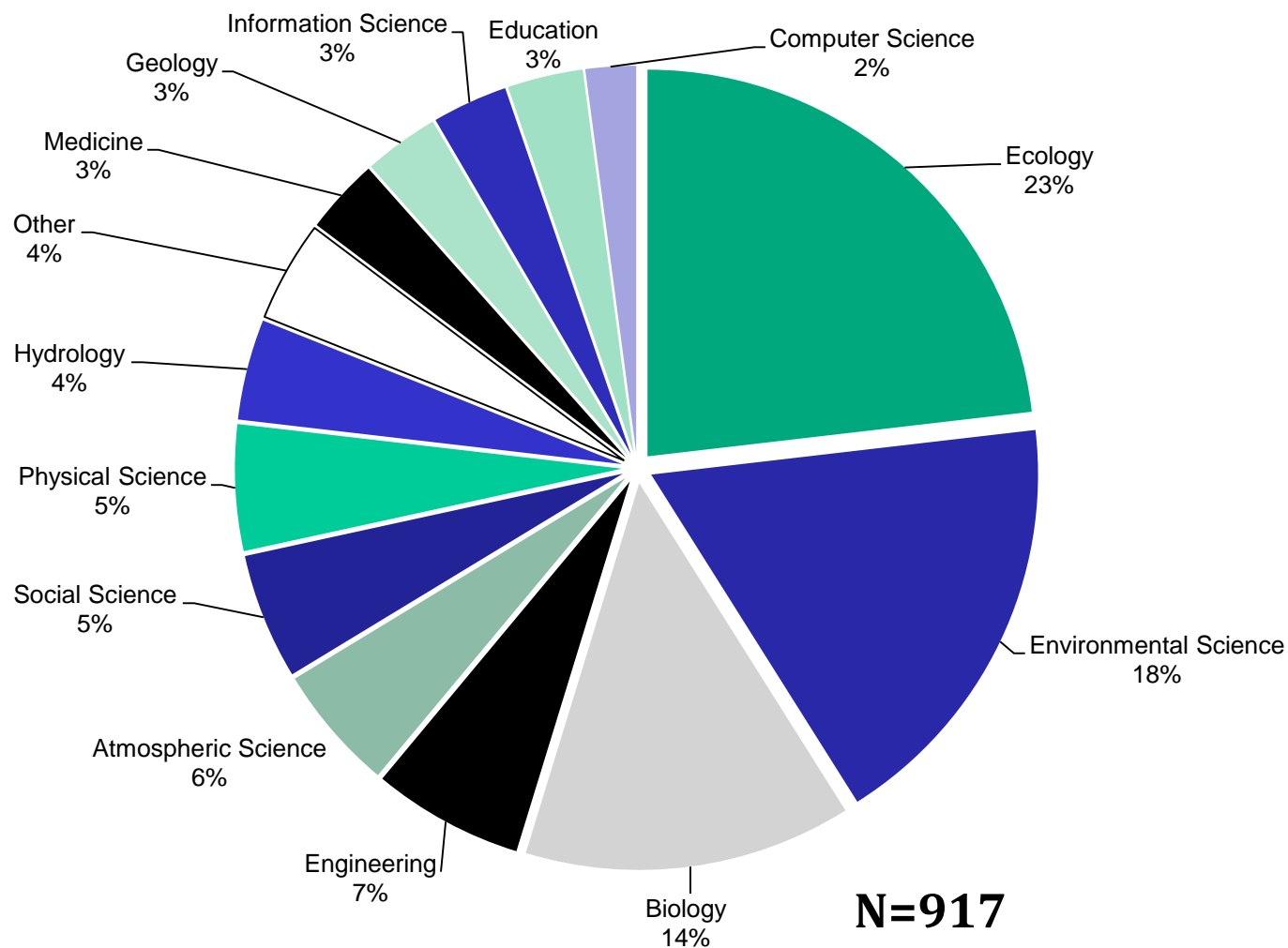
Primary Work Sector



N=917



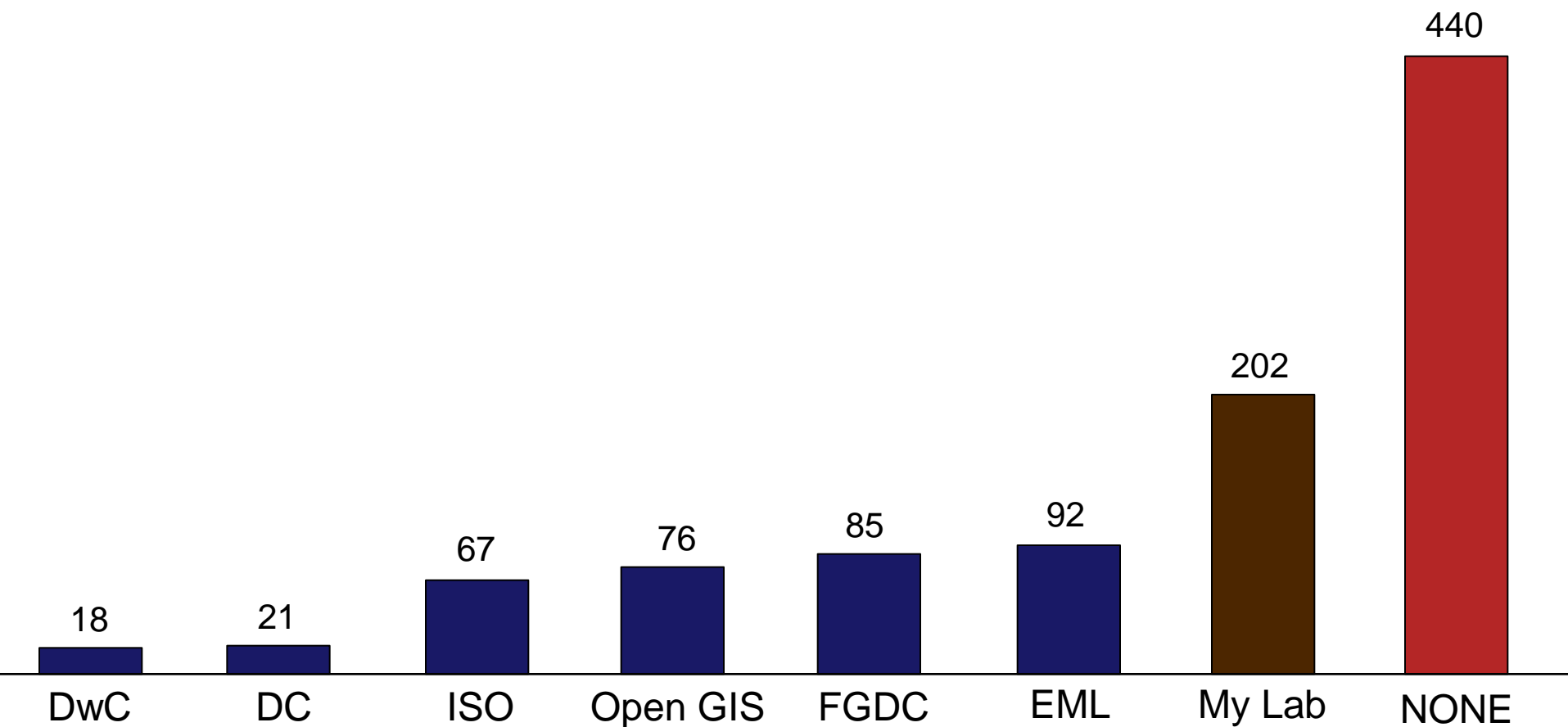
N=827



1. Data management **practices vary**.
2. Many scientists are **interested in sharing** data.
3. There are **many barriers** to sharing data.
4. There are some differences in **data management practices**.

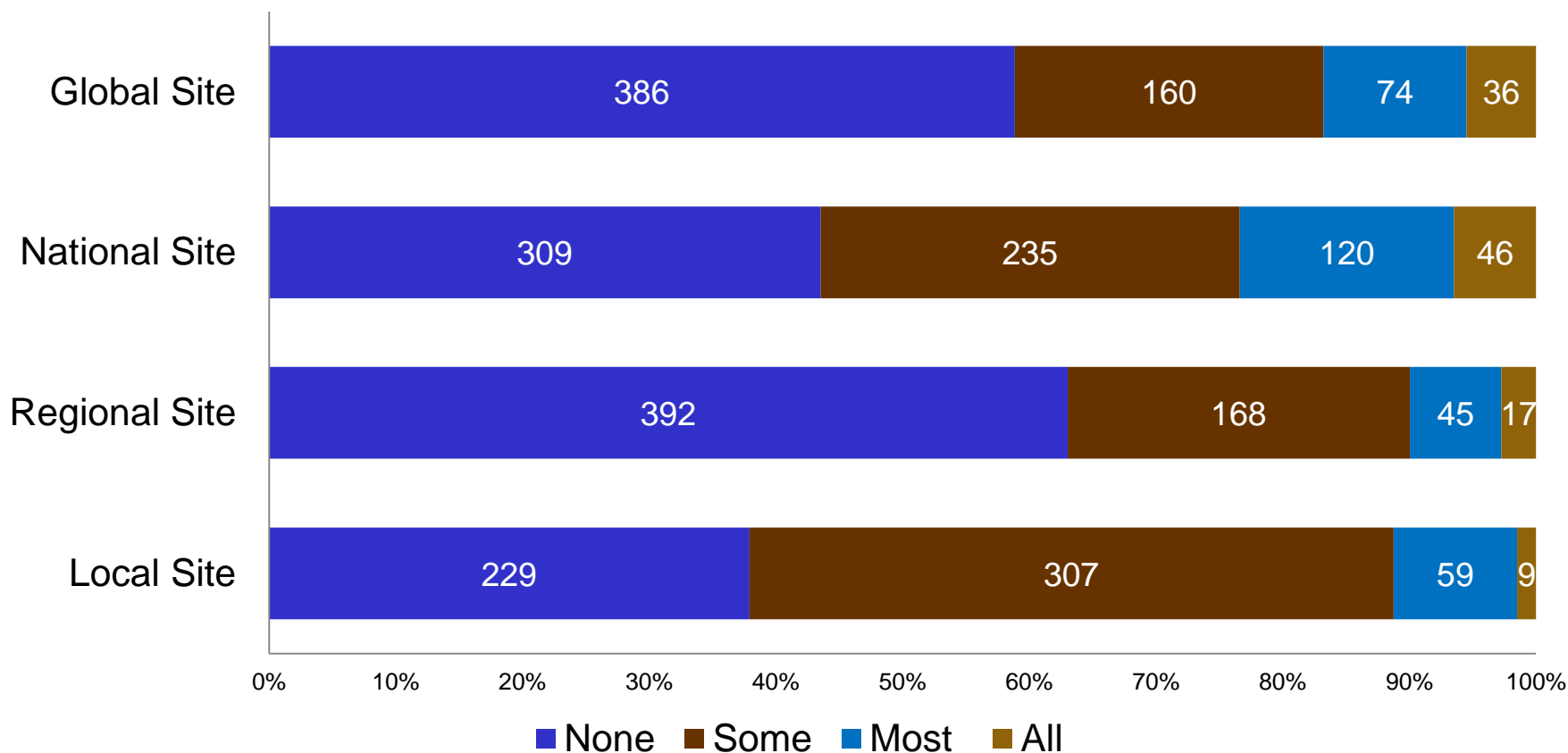
Data management practices vary.

What metadata do you currently use to describe your data, if any (check all that apply)?



- Data may be misinterpreted due to complexity of the data. (75%, N=899)
- Data may be misinterpreted due to poor quality of the data. (71%, N=899)
- Data may be used in other ways than intended. (74%, N=896)

If some or all of your data are available to others, these data are available:



Many scientists are interested
in sharing data.

- I would **use other researchers' datasets** if their datasets were easily accessible. (84%, N=902)
- I would be **willing to share data across a broad group** of researchers who use data in different ways. (83%, N=893)
- I would be willing to place at least **some of my data into a central data repository** with no restrictions. (79%, N=901)
- It is appropriate to **create new datasets from shared data**. (77%, N=902).
- I would be willing to place **all of my data into a central data repository** with no restrictions. (44%, N=894)

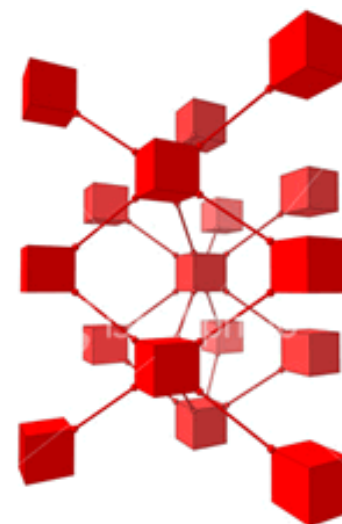
Condition	My Data	Others' Data
Acknowledge provider/funder	94%	94%
Formally cite provider/funder	94%	95%
Opportunity to collaborate	81%	82%
Reciprocal sharing agreement	71%	71%
Reprints of articles	70%	71%
Complete list of products	70%	69%

There are many
barriers to data sharing.

If your data are not available electronically to others, why not (check all that apply)?

- Insufficient time (54%)
- Lack of funding (41%)
- No place to put data (23%)
- Don't have the rights to make the data public (22%)

- Training on best practices (23%)
- Organization provides funds for long-term data management (24%)
- Organization provides funds for data management during project (31%)
- Others can access my data easily (38%)



Barriers for sharing research data (N=1270)

■ Legal Issues	41%
■ Misuse of data	41%
■ Incompatible data types	33%
■ Lack of Technical Infrastructure	28%
■ Lack of financial resources	27%
■ “Fear to lose” financial edge	27%
■ Restricted access to data archive	21%
■ No problems foreseen	16%
■ Other	10%

There are some differences in data management practices.

- Share data with others (78%)
- Others can access my data easily (50%)
- Org provides necessary tools during the project (58%)
- Org has process to manage data during the project (56%)
- Org provides storage beyond the project (54%)



- High satisfaction with data collection (82%, 73%)
- Data available on organization site (54%, 76%)
- Moderate satisfaction integrating data (43%, 44%)
- Tools to manage data during project (45%, 48%)
- Tools to store data beyond the project (38%, 54%)
- Low satisfaction with tools to prepare metadata (27%, 19%)

It is fair exchange for use of data when legal permission is obtained.

Age Range	My Data	Others' Data
30&Under	60%	62%
31-40	39%	39%
41-50	40%	42%
51-60	34%	35%
Over 60	32%	32%

At least part of the costs of data acquisition, retrieval, or provision must be recovered.

Age Range	My Data	Others' Data
30&Under	39%	42%
31-40	23%	25%
41-50	31%	33%
51-60	26%	26%
Over 60	30%	31%

Where do we go from here?



- Data management
- Identified many areas where D1 could learn from scientific communities
- Survey closes July 31, 2010
- Report in fall 2010